(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 April 2001 (26.04.2001)

PCT

(10) International Publication Number
WO 01/30090 A2

(51) International Patent Classification[7]: H04Q 7/00

(21) International Application Number: PCT/SE00/01906

(22) International Filing Date: 3 October 2000 (03.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/420,311 18 October 1999 (18.10.1999) US

(71) Applicant: TELEFONAKTIEBOLAGET LM ERICSSON (publ) [SE/SE]; S-126 25 Stockholm (SE).

(72) Inventor: SUGIRTHARAJ, David; Apartment #23, 4355 Grand Boulevard, Montreal, Quebec H4B 2X7 (CA).

(74) Agent: NORIN, Klas; Ericsson Radio Systems AB, Common Patent Department, S-164 80 Stockholm (SE).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/30090 A2

(54) Title: METHOD AND SYSTEM FOR THE WIRELESS TRANSMISSION OF LOSS SENSITIVE DATA

(57) Abstract: A method and system for improving the quality of service for data transmission to a mobile station (10). In accordance with one embodiment of the invention, a candidate cell (38(2)) for handoff of the mobile station is identified in response to the detection of a need to perform a handoff. Buffering resources for use in connection with the data transmissions are then allocated (82) in the candidate cell. In accordance with another embodiment, a condition causing a potential reduction in the quality of service to the mobile station is detected, and in response thereto, an existing allocation of buffering resources in the system is increased (56) to compensate for the potential reduction in the quality of service. In accordance with yet another embodiment of the invention, a condition causing a potential reduction in the quality of service to the mobile station is detected. In response thereto, a transmission rate of data to the mobile station is increased (116) to compensate for the potential reduction in the quality of service.

METHOD AND SYSTEM FOR THE WIRELESS TRANSMISSION OF LOSS SENSITIVE DATA

BACKGROUND OF THE INVENTION

Technical Field of the Invention

The present invention relates in general to wireless transmission of data, and in particular to the efficient processing of loss sensitive data services in a mobile telecommunications system.

Description of Related Art

Initially, mobile telecommunications systems were designed primarily for the purpose of enabling voice communications in a wireless environment. As wireless telecommunications systems have evolved, additional voice and messaging services and features have been added. In addition, newer generations of wireless telecommunications systems are being designed to handle an even more diverse range of data communications services. In contrast to analog and digital cellular protocols that are designed primarily for voice communications, these services, such as general packet radio service (GPRS), allow for the transmission of large amounts of data at high speeds. This capability makes it feasible to access the Internet and to retrieve applications software and data files using a wireless terminal.

Because of interference, multipath propagation, and other factors, however, the air interface that is used for communicating between a wireless mobile station and a fixed base station system is generally subject to greater data losses or bit errors than the hardwired connections that are used for transmitting packets of data via the Internet or other communications networks. Such losses are particularly prevalent during handoff of a mobile station from one base station to another because the handoff causes a temporary interruption in the connection. As a result, data losses and bit error rates tend to increase for a mobile station that is involved in frequent handoffs, such as when the mobile station is traveling from cell to cell at a rapid rate of speed.

If relatively small amounts of data are lost, or sporadic transmission errors occur, in connection with voice communications, a slight distortion in the transmitted

speech can result; such distortions typically do not present a substantial problem in understanding speech. In connection with the transmission of stacks of data, images, or software, on the other hand, the recipient is very sensitive to lost data and cannot afford to lose even a single byte of data. If a loss occurs, or if there are bit errors during transmission, such loss sensitive data must be re-transmitted.

To avoid having to retransmit data from its originating node, various intermediate nodes in a data communication network frequently include buffers that temporarily store packets of data, or a portion thereof, during the transmission process and for a short period after the packet is transmitted. Accordingly, when a loss or error occurs, the data can be re-transmitted from a nearby node rather than having to send a retransmission request all the way back to the originating node. If significant amounts of data have to be re-transmitted, however, the buffers can overflow, necessitating a retransmission of one or more data packets from the originating node. Such a retransmission from the originating node slows the effective rate of data transfer and reduces the overall quality of service that is provided to the mobile station.

One possible solution to this problem is to simply increase the buffer size. This solution, however, places additional demands on system resources in terms of processing and storage space, which increases the cost of the system and can cause further delays in transmission.

There is a need, therefore, for a method and system for minimizing the effective bit error rate and improving the quality of service in connection with the data transmissions in a wireless telecommunications system. Preferably, such a system and method would reduce the amount of retransmissions from an originating node while avoiding the need to design the system to incorporate large buffers for each different data connection in the wireless data network. In addition, such a method and system would include a process or mechanism for reducing the detrimental effects of handoff on data communications.

SUMMARY OF THE INVENTION

The present invention comprises a method and system for use in connection with data transmissions in a wireless communication system. In particular, the method

and system are used to improve the quality of service provided to a mobile station during the transmission of loss-sensitive data. When conditions in the wireless communication network are such that a loss of data or an increased bit error rate are likely to occur in connection with a particular data transmission session, the network reallocates resources to at least partially compensate for the poor transmission conditions.

In connection with one embodiment of the present invention, the system detects an impending need for handoff of the mobile station from a serving base station to a target base station during an ongoing data transmission. The system identifies the target base station as a candidate cell for the handoff at the mobile station in response to the detection of an impending need for a handoff. A signaling node that is used in the routing of data transmitted to the mobile station allocates buffering resources in the candidate cell prior to the handoff. These resources are allocated for use in connection with the data transmission. In addition, the signaling node begins transmitting data to the target base station prior to the handoff. This data can then be stored in the allocated buffering resources prior to completion of the handoff. The stored data is then available for transmission to the mobile station from the target base station immediately upon completion of the handoff, thereby reducing the likelihood of substantial data loss, and/or the detrimental effects of high bit error rates, during the handoff procedure.

In connection with an alternative embodiment of the present invention, a wireless data communication system includes a buffer for storing data to be transmitted to a mobile station. The buffer has a configurable amount of memory that is allocated for storing data in connection with data communications with the mobile station. Upon detection of a condition that might cause a reduction of quality of service to the mobile station, a signaling node that is used for routing data to the mobile station increases the amount of buffer memory that is allocated for the data communications. This increase serves to at least partially compensate for the potential reduction in the quality of service. Such a potential reduction in the quality of service can include, for example, a detection of a bit error rate that is above a selected threshold, a detection of a relatively high velocity of the mobile station, an

identification of a need for handoff of the mobile station, or a determination that a quantity of data stored in the allocated buffering resources exceeds the selected threshold. The system and method of this embodiment operate to help avoid the need to retransmit the data from an originating node by allowing increased amounts of the data to be stored in a node that is near the mobile station in the data transmission network.

In accordance with yet another embodiment of the present invention, the data communication system includes a server from which data to be transmitted to the mobile station originates. The system further includes a buffer in a wireless network that is used for storing data to be transmitted to the mobile station. Upon detection of a condition that might cause a reduction in the quality of service from the mobile station, a signaling node in the wireless network increases the transmission rate of data stored in the wireless network buffer to at least partially compensate for the potential reduction in the quality of service. Preferably, the increased transmission rate results from an allocation of additional transmission resources in a wireless network, such as an increase in the number of time slots used for transmitting data to the mobile station. As a result, data can be cleared from the wireless network buffer at a more rapid rate, thereby reducing the likelihood that data will need to be transmitted from the originating server during a period of increased data loss or of high transmission errors.

BRIEF DESCRIPTION OF THE DRAWINGS

For a more complete understanding of the present invention, reference is made to the following detailed description taken in conjunction with the accompanying drawings wherein:

FIGURE 1 is a block diagram of one type of data network for conducting wireless data communications;

FIGURE 2 is a graph showing an illustrative example of the quality of service to a mobile station;

FIGURE 3 shows a global positioning satellite system that can be used in connection with a mobile cellular network;

FIGURE 4 depicts a portion of the cellular network through which a highway passes;

FIGURE 5 illustrates a flow diagram of a method for improving a quality of a data communication service to a mobile station using dynamic buffer allocation in accordance with one embodiment of the invention;

FIGURE 6 shows a flow diagram of a method for improving the quality of a data communication service to a mobile station by early handoff initiation in accordance with another embodiment of the present invention;

FIGURE 7 illustrates a flow diagram of a method for improving the quality of a data communication service to a mobile station using early handoff initiation in accordance with an alternative embodiment of the present invention;

FIGURE 8 shows a flow diagram of a method for improving the quality of a data communication service to a mobile station using early handoff initiation in accordance with yet another alternative embodiment of the present invention;

FIGURE 9 illustrates a flow diagram of a method for improving the signal quality of a data communication service by increasing the transmission rate to a mobile station in accordance with another embodiment of the invention; and

FIGURE 10 depicts a flow diagram of a method for improving the signal quality of a data communication service by increasing the transmission rate to a mobile station in connection with a handoff in accordance with yet another embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

Reference is now made to the Drawings wherein like reference characters denote like or similar parts throughout the various Figures. Referring now to FIGURE 1, there is shown a block diagram of one type of data network for conducting wireless data communications. For purposes of illustration, the data network shown in the Figures and discussed below is a general packet radio service (GPRS) system 2. GPRS is a wireless data protocol that is used in connection with GSM standards. It will be understood by those of ordinary skill in the art, however, that the present

invention can also be used in connection with other types of packet switched networks and with circuit switched networks.

The types of data handled by a data network 2 can vary widely and might include, for example, data structures, images, and/or software. Most such data communications, as compared with voice communications, are particularly sensitive to data losses or bit errors. In fact, it is generally not acceptable for there to be any data loss or bit errors in connection with the transmission of loss sensitive data. It is, therefore, necessary to have mechanisms in place that permit correction of errors and retransmission of lost data. Such mechanisms are of particular importance in the wireless data communication context because the air interface that is used for wireless transmissions tends to be especially susceptible to errors or losses during transmission.

When data is requested by a mobile station 10 from an application server 12, the application server 12 retrieves the requested data and transmits a data packet containing the data, or a portion thereof, toward the requesting mobile station 10 via a packet network 16 such as the Internet. Generally, the packet network 16 includes a plurality of routing nodes 14. Each routing node 14 receives the transmitted data packet and retransmits it over the next hop to another routing node 14 in accordance with the address of the final destination (i.e., the mobile station 10).

Eventually, the data packet reaches a serving GPRS support node (SGSN) 18 that serves the area where the mobile station 10 is located. The SGSN 18 reformats the data packet for transmission in the GPRS system. This reformatting might involve, for example, converting the data packet from TCP/IP protocol to GPRS protocol. Once the data packet is reformatted, the SGSN 18 forwards the data packet to a serving base station (BS) 20 for the mobile station 10. The base station 20 in turn sends the data packet over an air interface 22 to the mobile station 10.

To facilitate data transmission, the SGSN 18 includes a buffer memory (M) 24 for temporarily storing data packets to be transmitted to the mobile station 10. Similarly, the base station 20 also includes a buffer memory (M) 26 for temporarily storing data packets to be transmitted to the mobile station 10. In cases where a transmission error occurs, data that is lost or data packets that contain bit errors upon receipt by the mobile station 10, as detected by an appropriate error detection protocol,

can be retransmitted to the mobile station 10 from one of the buffer memories 24 or 26. This ability to retransmit stored data from a relatively local node avoids the delay that would be caused by having to request and perform a retransmission from the application server 12.

Once the data is successfully received by the mobile station 10, a confirmation message can be sent to the base station 20 and the SGSN 18 providing a notification of the successful transmission. The data in the buffer memories 24 and 26 can then be deleted, thereby freeing up room for additional data packets to be received from the packet network 16 for transmission to the mobile station 10.

If, however, relatively large amounts of data are lost or frequent bit errors occur, the amount of buffer space allocated for data communications with the mobile station 10 can overflow. In other words, data packets to be transmitted to the mobile station 10 might be received by the SGSN 18 and/or the base station 20 at a faster rate than data is being deleted from the buffer memory 24 or 26. Such an occurrence causes some data to be lost because there is not sufficient buffer space to store all of the data. The loss of data necessitates retransmission of data from the application server 12, thereby reducing the quality of service to the mobile station 10 by slowing the rate of data transfer.

One of the primary causes of data loss or errors in transmission is handoff of a mobile station 10 from one base station 20 to another. As the mobile station 10 moves toward the edge of a cell served by the currently serving base station 20, the signal strength of signals received by the mobile station 10 tends to decrease. In a typical cellular system, the mobile station 10 periodically measures the strength of signals from the serving base station and from base stations 20 in adjacent cells. The mobile station 10 forwards this information to the serving base station 20 so that a determination can be made as to when handoff to another base station 20 is necessary. Once the signal strength from the serving base station 20 falls below a predetermined threshold, the wireless network recognizes that a handoff should be executed. Based on the measured strengths of signals from base stations 20 in adjacent cells, the wireless network identifies a target cell and initiates a handoff (as indicated at 28) of

the mobile station 10 from the serving base station 20 to a target base station 20' in the target cell.

Once the handoff is complete, communications with the mobile station 10 are conducted via an air interface 22' between the mobile station 10 and the target base station 20', which becomes the new serving base station 20'. Data to be transmitted to the mobile station 10 is then sent by the SGSN 18 to the new serving base station 20' so that it can be transmitted to the mobile station 10. During handoff, however, communications with the mobile station 10 are temporarily interrupted. This interruption will frequently result in some data being lost, thereby necessitating a retransmission of the lost data. In some cases, the loss of data can even cause the buffer memory 24 or 26 to overflow, which necessitates retransmission from the application server 12.

When the velocity of the mobile station 10 increases, the mobile station 10 travels across the cells in the wireless network more rapidly, causing more frequent handoffs between cells. As the frequency of handoffs increases, problems with lost data and bit errors tend to multiply, further reducing the effective rate of data transfer. In addition, frequent handoffs can place more demands on network processing resources (e.g., logic processing and data storage), which can also affect the data transfer rate.

Furthermore, reductions in the rate of data transfer can occur because data transmissions to the mobile station 10 are not possible under certain radio conditions. When the mobile station 10 is unable to receive data, a build-up and overflow of data in the buffers 24 and 26 can occur. For example, when the mobile station 10 is near the serving base station 20 in a cell, signal quality and signal strength for signals received by the mobile station 10 are normally very high. Such high signal quality and signal strength result in relatively low data loss and a low bit error rate. As the distance between the serving base station 20 and the mobile station 10 increases, however, received signal quality and received signal strength tend to deteriorate, which causes an increase in the amount of data that is lost and/or an increase in the bit error rate.

Referring now to FIGURE 2, there is depicted a graph showing an illustrative example of the quality of service provided to a mobile station 10 versus the proximity of the mobile station 10 to the base station 20. When the mobile station 10 is very near the base station 20, the quality of service provided (as indicated by line 40) tends to be at or near its maximum (as indicated at 42). On the other hand, when the mobile station 10 approaches the edge of the coverage area for a particular cell, the quality of service 40 provided is substantially reduced (as indicated at 44). As the quality of service 40 deteriorates, a bit error rate (BER) 46 for communications with the mobile station 10 correspondingly increases. Ideally, a network operator could implement procedures or mechanisms in the network to provide a more consistent effective quality of service 40 to the mobile station so as to reduce the bit error rate and/or data loss as the mobile station 10 nears the edge of the serving cell, or as the mobile station 10 becomes involved in a handoff.

In accordance with the present invention, there are several different methods and systems that can be used to improve the quality of service provided to a mobile station 10 in a wireless data network. In a first embodiment, the size of the buffers 24 and 26 used to store data packets in the wireless data communication system can be dynamically increased as necessary to compensate for substantial data losses or high bit error rates. Current GPRS systems have static buffers; once the buffers reach capacity, the buffer begins to overflow, causing a loss of data. By dynamically increasing the buffer size to meet storage/transmit needs, however, the need to retransmit data from the application server 12 can be substantially reduced.

In accordance with another embodiment of the invention, data is forwarded to a target base station 20' prior to the completion of a handoff. Existing wireless data communication systems generally use a concept of capacity on demand. Thus, resources in the target cell are not allocated for use in data communications until the mobile station 10 is actually in communication with the target base station 20' (i.e., when handoff is substantially complete). If data is lost during handoff, that data must be retransmitted. By forwarding data to the target base station 20' prior to handoff, on the other hand, data is immediately available for transmission to the mobile station 10 upon completion of handoff, thereby improving the quality of service provided to the

mobile station 10. To facilitate the early forwarding of data, the handoff process can also be initiated earlier to further reduce the probability that significant amounts of data might be lost.

In another embodiment of the invention, handoff of the mobile station 10 can be initiated earlier and more efficiently by predicting an impending handoff based on a calculated velocity and direction of travel of the mobile station 10. As the mobile station 10 travels through the data network 2, the network 2 monitors the position of the mobile station 10. Based on detected changes in the mobile station's position over time, the velocity and direction of travel can be calculated. An approximate time and target cell for handoff can then be predicted using the velocity vector information in connection with a known cellular network topology. Using these predictions, handoff can be initiated earlier and accomplished more quickly, thereby reducing the potential for significant data loss.

In accordance with yet another embodiment of the present invention, the transmission rate of data to the mobile station 10 can be increased to compensate for actual or anticipated data losses or high bit error rates. In a typical GPRS system, a subscriber negotiates a certain level of service. For example, depending on how much the subscriber is willing to pay, the subscriber might be able to receive between approximately 14 kilobits per second (kbs), which corresponds to an allocation of one time slot to the data transmission, and approximately 115 kbs, which corresponds to an allocation of eight time slots. The level of service is then determinative of the maximum number of time slots that will be allocated to the subscriber at any one time. In addition to preventing time slots in a target cell from being allocated while the mobile station 10 is still operating in another cell, this limit also prevents the mobile station 10 from being able to receive data at a faster rate than what has been negotiated by the subscriber. By allocating additional time slots to the mobile station, however, when the bit error rate is high, buffers can be emptied quicker, and the need for retransmissions can be avoided.

Referring now to FIGURE 3, there is illustrated a global positioning satellite (GPS) system 4 that can be used in connection with a mobile cellular network 2 to monitor the position of a mobile station 10. The GPS system 4 includes a plurality of

positioning satellites 30 that continuously transmit signals (as indicated at 31) containing satellite location and time reference data. Incorporated into the mobile station 10 is a GPS receiver 32 designed to receive the signals transmitted by the positioning satellites 30. Using the data contained in the signals from three or more different positioning satellites 30, the GPS receiver 32 can determine a current location, velocity vector, and time reference for the mobile station 10 in accordance with known processing methods. This information can be displayed at the mobile station 10 and can further be forwarded to a serving base station 20 over the air interface 22 so that the cellular network 2 can monitor the mobile station's position, velocity, and direction of travel.

Referring now to FIGURE 4, there is illustrated a portion of the cellular network 2 through which a highway 34 passes. As a mobile station 10 travels along the highway 34 at a relatively high velocity, it moves from one cell 38 to the next at a rate that is proportional to its velocity. As a result, handoffs must be performed more frequently than in the case of a mobile station 10 that is slow moving or stationary. Particularly in the case of data transmissions, these frequent handoffs can substantially reduce the quality of service to the mobile station 10. Furthermore, in the case of high velocity mobile stations 10, there is a smaller time window in which to perform handoff as compared with slower moving mobile stations 10.

As discussed in connection with FIGURE 3, a GPS receiver 32 included in the mobile station 10 can be used to inform the cellular network 2 of the position, velocity, and direction of travel for the mobile station 10. Alternatively, other procedures for locating the mobile station 10 can be used. For example, the mobile station's position can be triangulated using a measured propagation delay of signals between the mobile station 10 and multiple base stations 20. As the mobile station 10 travels in the direction indicated by arrow 33, it approaches a cellular boundary 36 between a serving cell 38(1) served by a serving base station 20 and an adjacent cell 38(2) served by a target base station 20'. Using the velocity vector information, the network 2 can predict the approximate time at which handoff will be necessary. In addition, the network 2 can reduce the list of candidate cells 38 for handoff so as to predict which cell or cells 38 are most appropriate for handoff based on the detected direction of

travel 33. This allows the network 2 to begin the handoff process earlier (i.e., before the mobile station 10 reaches the cell boundary 36) and to conduct handoff more efficiently, thereby reducing the probability or extent of data loss during handoff.

Referring now to FIGURE 5, there is illustrated a flow diagram 50 of a method for improving a quality of a data communication service to a mobile station 10 using dynamic buffer allocation in accordance with one embodiment of the invention. At step 52, a data transmission is initiated, wherein loss sensitive data is retrieved from an application server 12 and sent to the mobile station 10 via a GPRS network 2. During the data transmission, the network detects, through the use of a GPS system 4 or other positioning system, that the mobile station 10 is traveling at a high velocity at step 54. To avoid a potential reduction in the quality of service (e.g., through data loss or an increase in the bit error rate), which can result from the high velocity of the mobile station 10, the amount of buffer space allocated to the mobile station 10 in the buffer memories 24 and 26 of the SGSN 18 and the base station 20 is increased at step 56. This increase in buffer space helps prevent buffer overflow and permits data retransmissions, if necessary, to be made from a relatively local node rather than from the distant application server 12.

Subsequently, the velocity of the mobile station 10 is monitored to determine whether it has decreased at step 58. Also at step 58, the data transmission is monitored to determine if it is complete. If neither of these conditions is satisfied, the allocated buffer space remains the same and the monitoring continues. However, if the velocity of the mobile station 10 decreases or if the data transmission is complete, the amount of buffer space allocated to the mobile station 10 is reduced back to normal levels at step 60. Although the method illustrated in FIGURE 5 is shown and described in connection with a high velocity mobile station 10, it is also possible to implement the procedure in connection with other circumstances involving a reduction in the quality of service to the mobile station 10. For example, buffer space increases can also be used in cases where high bit error rates are detected, where the system 2 or mobile station 10 determines that a handoff is necessary, or when the buffer begins to approach capacity. By reducing the amount of data loss through increased buffer allocations, fewer re-transmissions are required and quality of service is improved.

Referring now to FIGURE 6, there is illustrated a flow diagram 70 of a method for improving the quality of a data communication service to a mobile station 10 by early handoff initiation in accordance with another embodiment of the present invention. At step 72, a data transmission is initiated, wherein loss sensitive data is retrieved from an application server 12 and sent to the mobile station 10 via a GPRS network 2. During the transmission, the network 2 detects that the mobile station 10 is traveling at a high velocity at step 74. As a result, the signal threshold for handoff is increased at step 76, causing handoff to be initiated earlier. In particular, as the mobile station 10 moves toward the edge of the serving cell, the signal strength begins to decrease. Once the signal strength falls below the signal threshold, the mobile station 10 and/or the cellular network begins to search for candidate cells for handoff. Because the signal threshold for the high velocity mobile station 10 has been increased, however, this initiation of the handoff procedure begins earlier than in a standard handoff scenario.

At step 78, the network 2 detects an impending handoff (i.e., the detected signal strength falls below the signal threshold). Next, at step 80, the network 2 identifies one or more candidate cells for handoff of the mobile station 10. Preferably, the network 2 uses the velocity vector for the mobile station 10 to predict which cell or cells are the most likely cells for handoff and uses these predictions to reduce the list of candidate cells. Once a single or small number of candidate cells are identified, the network 2 allocates resources in the candidate cell or cells at step 82. Preferably, this resource allocation includes an allocation of buffer space in the candidate cell. This allows data to be forwarded to and stored in the candidate cell in advance of handoff, so that data transmissions can resume immediately upon completion of handoff. By providing longer notice of handoff, resource demands can be spread over time, the risk of losing calls during handoff is reduced, and there is a lower risk of losing data during handoff between cells. In addition, if data is lost during handoff, the re-transmission of the lost data can be handled by the candidate cell using the buffered data. This results in a much quicker re-transmission of lost data. At some point after the resource allocation in the candidate cell, handoff is performed at step 84.

Referring now to FIGURE 7, there is illustrated a flow diagram 90 of a method for improving the quality of a data communication service to a mobile station 10 using early handoff initiation in accordance with an alternative embodiment of the present invention. At step 72, a data transmission is initiated. Next, the network 2 predicts the approximate dwell time of the mobile station 10 in the serving cell at step 92. The dwell time can be estimated using a detected velocity vector for the mobile station 10 in conjunction with a known cellular topology of the cellular system 2. Alternatively, the dwell time can be predicted based on recent dwell times for the mobile station 10 in previous cells. In any event, use of a predicted dwell time is primarily applicable during highway travel, where velocity is relatively constant and the direction of travel can be predicted with a relatively high degree of accuracy.

Using the predicted dwell time, an impending handoff is identified at step 78. Subsequently, the process continues in accordance with steps 80-84, as shown in and discussed in connection with FIGURE 6. For example, if the network determines that the dwell time in the serving cell for the mobile station 10 is twenty seconds, the network might allocate buffer space in the candidate cell and begin storing data there after the mobile station 10 has been in the serving cell for ten seconds.

Referring now to FIGURE 8, there is illustrated a flow diagram 100 of a method for improving the quality of a data communication service to a mobile station 10 using early handoff initiation in accordance with yet another alternative embodiment of the present invention. At step 72, a data transmission is initiated. Then, at step 102, the network 2 detects a diminishing signal quality of signals between the serving base station 20 and the mobile station 10. As illustrated in FIGURE 2, such a diminished signal quality is frequently caused by the mobile station 10 moving away from the base station 20 in the serving cell. In response to the detected diminishing signal quality, the network 2 identifies a candidate cell for handoff at step 80, allocates resources in the candidate cell at step 82, and performs handoff at step 84, as in the methods shown in FIGURES 6 and 7.

Referring now to FIGURE 9, there is illustrated a flow diagram 110 of a method for improving the signal quality of a data communication service by increasing the transmission rate to a mobile station 10 in accordance with another embodiment

of the invention. At step 112, a transmission of data to the mobile station 10 is initiated. As discussed above, such a data transmission generally involves a predetermined maximum number of time slots, depending upon the level of service negotiated by the subscriber. This maximum number of time slots limits the rate at which data can be transmitted to the mobile station 10. In accordance with the invention, however, the network 2 detects a high bit error rate for communications with the mobile station 10 at step 114. In response, the network 2 temporarily increases the transmission rate for communications with the mobile station 10 by allocating additional time slots, assuming that they are available in the serving cell, to the mobile station 10 at step 116. This increased transmission rate effectively prevents a noticeable decline in the quality of service to the mobile station 10. Although FIGURE 9 depicts increasing the transmission rate in connection with a high bit error rate, this procedure can also be used to improve the quality of service in other situations. For example, the transmission rate can also be increased when the buffer memory begins to approach capacity, thereby emptying data from the buffer at a faster rate.

Referring now to FIGURE 10, there is illustrated a flow diagram 120 of a method for improving the signal quality of a data communication service by increasing the transmission rate to a mobile station 10 in connection with a handoff in accordance with yet another embodiment of the present invention. At step 112, a data transmission is initiated. Subsequently, at step 78, the network 2 determines that a handoff is imminent (e.g., in accordance with the procedures discussed in connection with FIGURES 6-8). The network 2 then identifies a candidate cell for handoff at step 80, allocates resources in the candidate cell at step 82, and performs handoff at step 84.

By allocating buffer space in the candidate cell and storing data there (at step 82), a significant amount of data might be stored in the candidate cell upon completion of handoff. Therefore, the transmission rate in the candidate cell is temporarily increased at step 116 so that the stored data can be at least partially cleared from the buffer. Once the buffer has been cleared, the transmission rate is returned to normal at step 118. This decrease in the transmission rate can occur, for example, once the

buffer has been emptied below a selected threshold or after a certain number of data packets have been transmitted in the new cell.

As shown in FIGURE 10, early handoff procedures can be used in conjunction with increases in the transmission rate to improve service quality. Similarly, dynamic buffering procedures can also be used in connection with early handoff procedures and/or increases in the transmission rate to improve service quality.

Although a preferred embodiment of the method and apparatus of the present invention has been illustrated in the accompanying Drawings and described in the foregoing Detailed Description, it is understood that the invention is not limited to the embodiment disclosed, but is capable of numerous rearrangements, modifications, and substitutions without departing from the spirit of the invention as set forth and defined by the following claims.

WHAT IS CLAIMED IS:

1. A method for improving the quality of service for a transmission of loss sensitive data to a mobile station, comprising the steps of:

detecting an impending need for handoff of the mobile station receiving a data transmission;

identifying a candidate cell for handoff of the mobile station in response to the detection of the impending need for handoff; and

allocating buffering resources associated with the candidate cell prior to handoff, said resources for use in connection with the data transmission.

2. The method of claim 1, further comprising the step of sending data to be transmitted to the mobile station to the candidate cell, said data sent to the candidate cell prior to handoff.

3. The method of claim 2, further comprising the step of storing the data to be transmitted to the mobile station in the candidate cell.

4. The method of claim 3, further comprising the steps of:

performing a handoff of the mobile station to the candidate cell; and

transmitting the stored data to the mobile station in the candidate cell.

5. The method of claim 4, further comprising the step of increasing a transmission rate of data transmitted to the mobile station after performing the handoff.

6. The method of claim 1, further comprising the step of monitoring a velocity of the mobile station, wherein the step of detecting the impending need for handoff including predicting an approximate time of handoff based on the monitored velocity.

7. The method of claim 6, further comprising the step of raising a signal threshold for initiating handoff in response to a monitored velocity of the mobile station.

8. The method of claim 6, wherein the prediction of the approximate time of handoff includes predicting an approximate dwell time in a serving cell of the mobile station.

9. The method of claim 1, wherein the step of detecting an impending need for handoff includes detecting a diminishing signal quality of a communication with the mobile station.

10. The method of claim 1, further comprising the step of monitoring a velocity vector of the mobile station, said velocity vector used to reduce a list of potential candidate cells for handoff of the mobile station.

11. A method for improving a quality of service to a mobile station receiving a data transmission in a wireless data communication system, comprising the steps of:

detecting a condition causing a potential reduction in the quality of service to the mobile station;

dynamically increasing an allocated amount of buffering resources in the wireless data communication system in response to the detected condition, said dynamically increased allocation for use in connection with the data transmission to at least partially compensate for the potential reduction in the quality of service.

12. The method of claim 11, wherein the step of detecting said condition comprises detecting a bit error rate above a selected threshold bit error rate.

13. The method of claim 11, wherein the step of detecting said condition comprises detecting a velocity of the mobile station, said velocity exceeding a selected threshold velocity.

14. The method of claim 11, wherein the step of detecting said condition comprises identifying a need for handoff of the mobile station.

15. The method of claim 11, wherein the step of detecting said condition comprises determining that a quantity of data stored in the allocated amount of buffering resources exceeds a selected threshold quantity of data.

16. The method of claim 11, further comprising the step of storing data to be transmitted to the mobile station in the increased allocation of buffering resources.

17. The method of claim 16, further comprising the steps of:

transmitting data stored in the increased allocation of buffering resources;

detecting a bit error rate for the transmitted data, said bit error rate exceeding a selected threshold bit error rate; and

re-transmitting the data stored in the increased allocation of buffering resources in response to the detected bit error rate.

18. The method of claim 16, further comprising the steps of:

transmitting data stored in the increased allocation of buffering resources;

detecting a loss of the transmitted data; and

re-transmitting the data stored in the increased allocation of buffering resources in response to the detected data loss.

19. A method for improving a quality of service to a mobile station receiving a data transmission in a wireless data communication system, comprising the steps of:

detecting a condition causing a potential reduction in the quality of service to the mobile station; and

increasing a transmission rate of data to the mobile station in response to the detected condition, said increased transmission rate to at least partially compensate for the potential reduction in the quality of service.

20. The method of claim 19 wherein the step of increasing the transmission rate includes allocating additional transmission resources in the wireless data communication system for the data transmission.

21. The method of claim 20, wherein the additional transmission resources comprise at least one additional time slot for transmitting data to the mobile station.

22. The method of claim 19, wherein the step of detecting said condition comprises detecting a bit error rate above a selected threshold bit error rate.

23. The method of claim 19, wherein the step of detecting said condition comprises identifying a need for handoff of the mobile station.

24. The method of claim 19, wherein the step of detecting said condition comprises determining that a quantity of data stored in an allocated amount of buffering resources exceeds a selected threshold quantity of data.

25. A wireless data communication system, comprising:

a serving base station transmitting data to a mobile station; and

a target base station selected as a candidate for handoff of the mobile station; and

a signaling node for routing the data transmitted to the mobile station, said signaling node routing data to be transmitted to the mobile station to the target base station prior to handoff as a result of a predicted mobile station handoff to the target base station.

26. The system of claim 25, wherein the predicted handoff is identified in accordance with a detected velocity of the mobile station.

27. The system of claim 26, wherein the velocity of the mobile station is detected by a global positioning satellite system receiver associated with the mobile station.

28. The system of claim 26, wherein the velocity of the mobile station is detected using a measured propagation delay of signals between the mobile station and a plurality of base stations.

29. The system of claim 25, wherein the target base station is selected in accordance with a detected direction of travel of the mobile station.

30. The system of claim 25, further comprising a buffer associated with the target base station for storing the data to be transmitted to the mobile station.

31. The system of claim 25, wherein the wireless data communication system comprises a circuit switched network.

32. The system of claim 25, wherein the wireless data communication system comprises a packet switched network.

33. The system of claim 32, wherein the packet switched network comprises a general packet radio service (GPRS) system.

34. The system of claim 33, wherein the signaling node comprises a serving GPRS support node.

35. A wireless data communication system, comprising:

a buffer for storing data to be transmitted to a mobile station, said buffer having a configurable amount of memory allocated for storing data in connection with data communications with the mobile station; and

a signaling node for routing to the mobile station the data to be transmitted, said signaling node dynamically increasing the amount of memory allocated for the data communications in response to a detection of a potential reduction in the quality of service to the mobile station.

36. The system of claim 35, wherein the buffer is associated with a base station, said base station for transmitting data to the mobile station over an air interface.

37. The system of claim 35, wherein the buffer is associated with the signaling node.

38. The system of claim 35, wherein the signaling node comprises a serving general packet radio service support node.

39. The system of claim 35, wherein the signaling node comprises a base station for transmitting data to the mobile station over an air interface.

40. A data communication system, comprising:

a server sending data to be transmitted to a mobile station, said data to be transmitted via a wireless network;

a wireless network buffer for storing data to be transmitted to a mobile station; and

a wireless network signaling node for routing to the mobile station the data to be transmitted, said signaling node increasing a transmission rate of data stored in the wireless network buffer in response to a detection of a potential reduction in the quality of service to the mobile station.

41. The system of claim 40, wherein the wireless network buffer is associated with a base station for transmitting data to the mobile station over an air interface.

42. The system of claim 40, wherein the signaling node increases the transmission rate of data by assigning additional slots for transmitting the data to be transmitted.

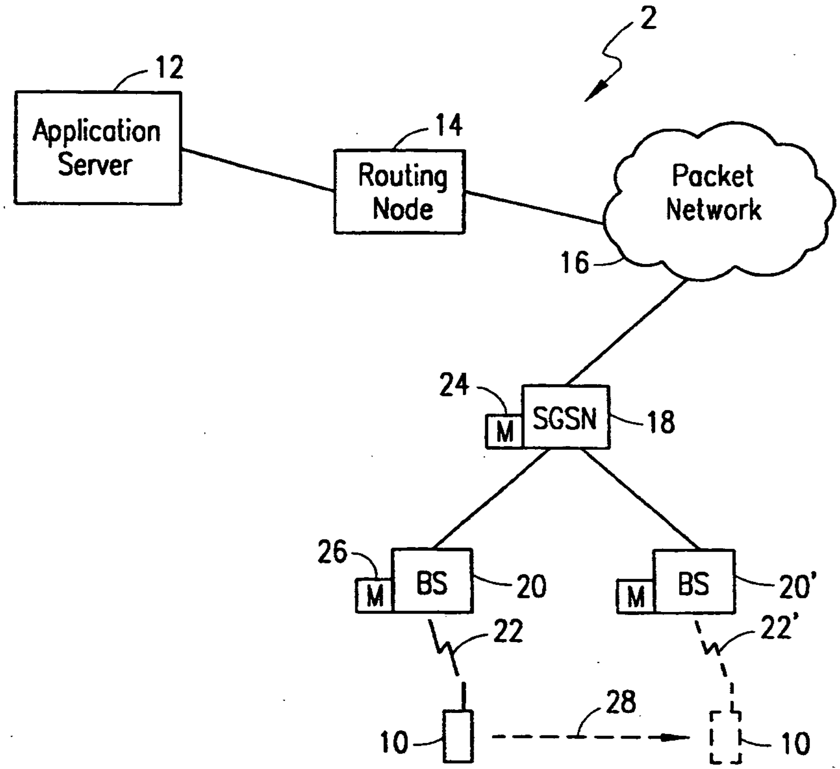
2
12
Application Server
14
Routing Node
Packet Network
16
24
M
SGSN
18
26
M
BS
20
M
BS
20'
22
22'
28
10
10

FIG. 1

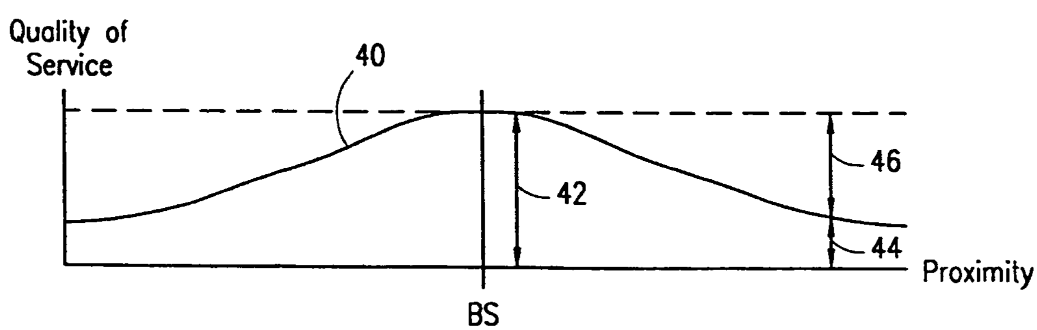
Quality of Service
40
46
42
44
Proximity
BS

FIG. 2

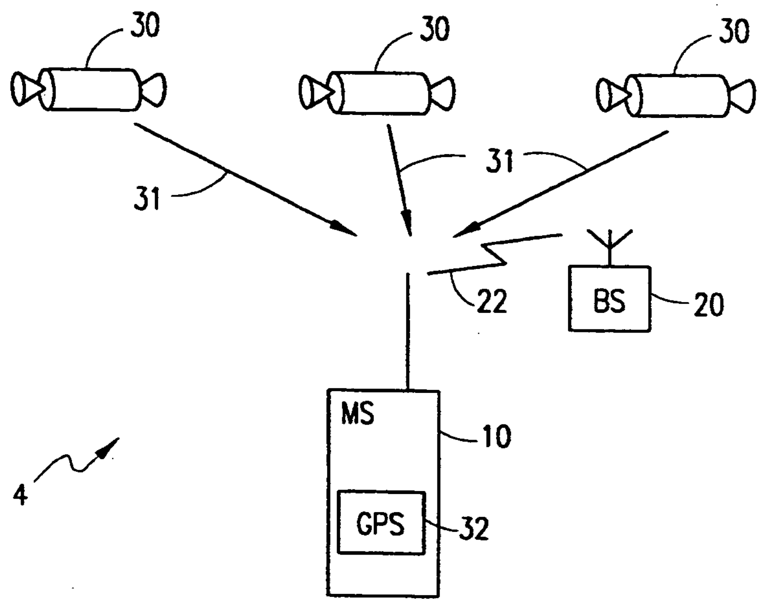
30
30
30
31
31
22
BS
20
MS
10
4
GPS
32

FIG. 3

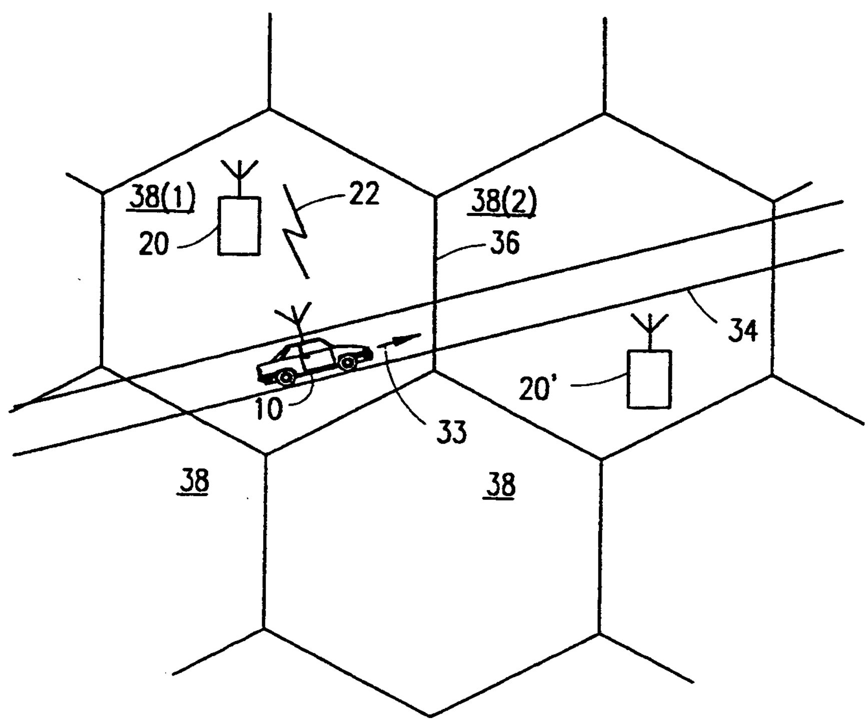
38(1)
22
38(2)
20
36
34
20'
10
33
38
38

FIG. 2

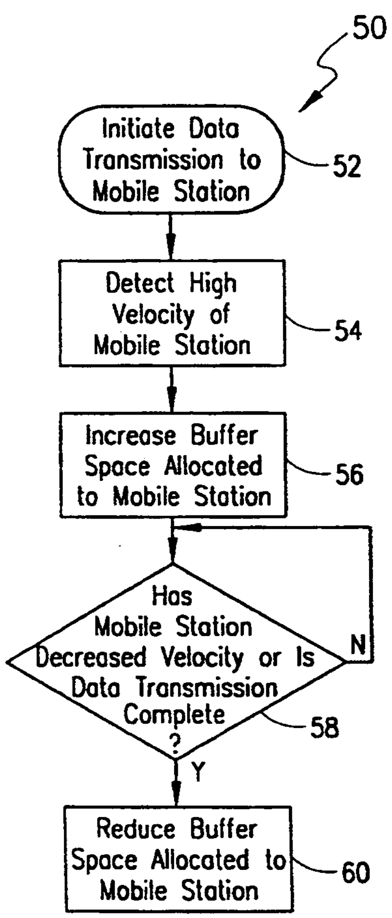

FIG. 5

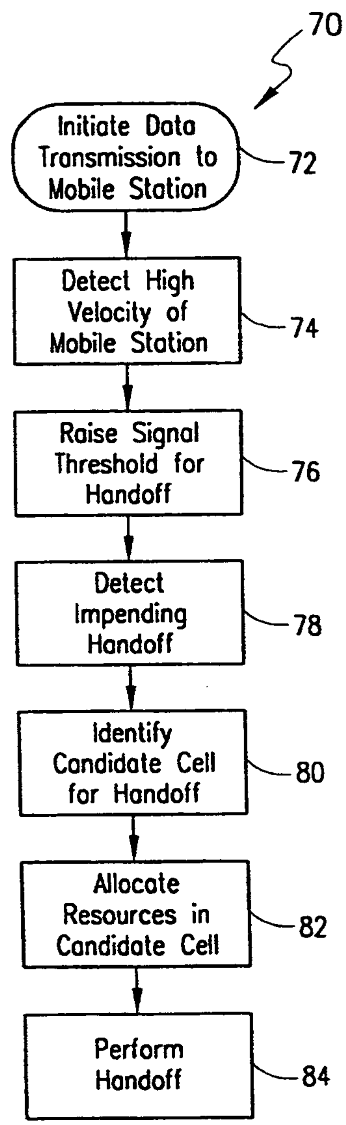

FIG. 6

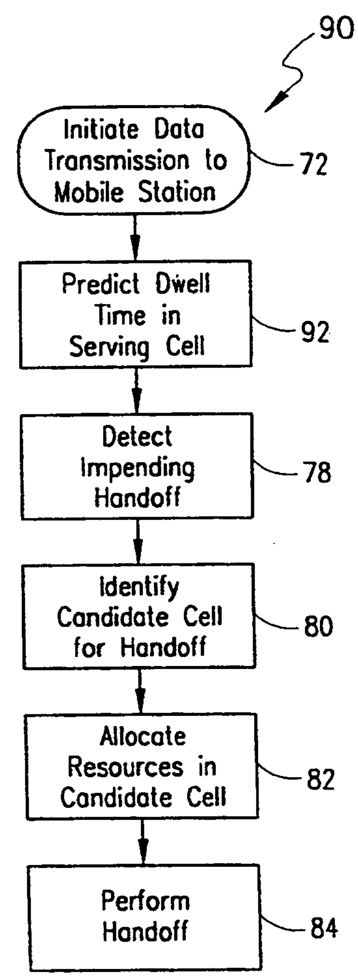

FIG. 7

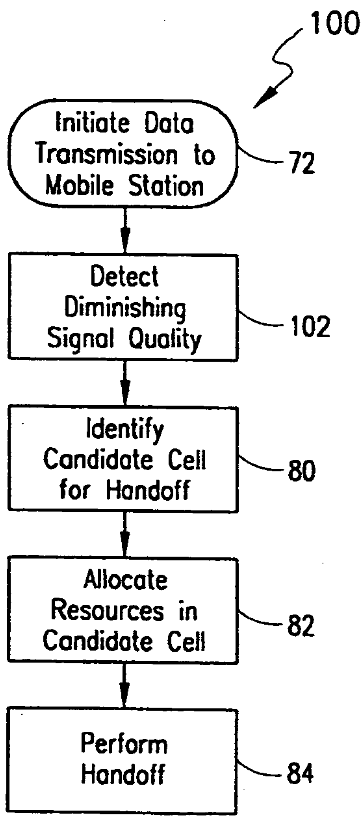

FIG. 8

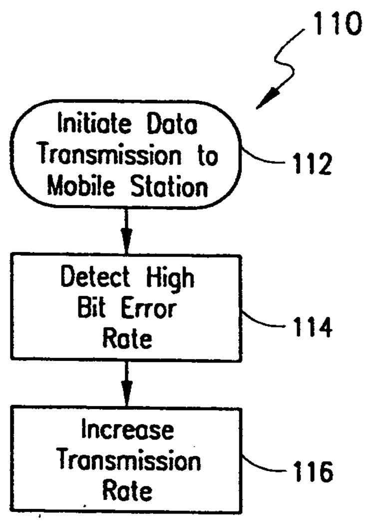

FIG. 9

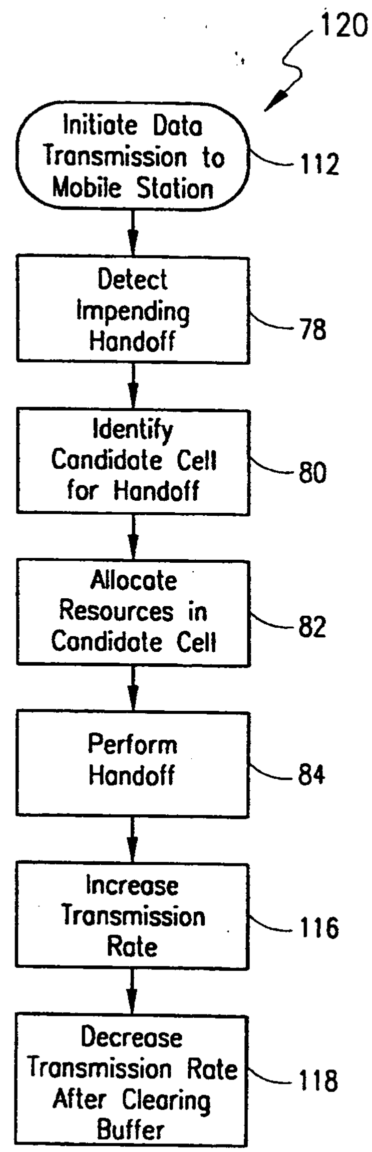

FIG. 10